Gödel's Incompleteness Theorems

by Sally Cockburn (2016)

1. Gödel Numbering

We begin with Peano's axioms for the arithmetic of the natural numbers (*i.e.* number theory):

(1) Zero is a natural number.

(2) Every natural number has an immediate successor that is also a natural number.

(3) Zero is not the immediate successor of any natural number.

(4) If two natural numbers have equal immediate successors, they are themselves equal.

(5) If a set of natural numbers A contains zero as well as the immediate successor of any natural number in A, then every natural number is in A.

Assume that S is a formal language powerful enough to represent Peano arithmetic. For the sake of simplicity, we assume that S is a somewhat richer language than ZF in that it includes symbols for the unary immediate successor operation and the binary operations of addition and multiplication. There are many different ways of assigning numbers to the components of the formal language S; we give only one example. We begin by assigning numeric symbols to the symbols according to Table 1.

Table 1. Numeric Symbols

symbol	numeric symbol	meaning
$\neg$	1	not
$\wedge$	3	and
$\vee$	5	or
$\Longrightarrow$	7	implies
$\Longleftrightarrow$	9	if and only if
$\forall$	11	for all
$\exists$	13	there exists
$=$	15	equals
(	17	left parenthesis
)	19	right parenthesis
0	21	zero
s	23	immediate successor
$+$	25	plus
$\times$	27	times

We also want a system of numeric symbols for variables, representing either numbers or well-formed formulas. For this, we use powers of primes greater than or equal to 29. The numeric symbol of a numerical variable is a prime, while the numeric symbol of a variable representing a well-formed formula with k free variables is a prime raised to the power $k+2$.

TABLE 2. Numeric Symbols for Variables

symbol	number	meaning
x	29	numerical variable
y	31	numerical variable
$\vdots$	$\vdots$	$\vdots$
p	29^2	propositional variable
q	31^2	propositional variable
$\vdots$	$\vdots$	$\vdots$
ϕ	29^3	predicate variable
ψ	31^3	predicate variable
$\vdots$	$\vdots$	$\vdots$

If p is a string of n symbols, with corresponding numeric symbols $m_1, m_2, \ldots, m_n$, then we define the Gödel number of p to be

$$GN(p) = \pi_1^{m_1} \cdot \pi_2^{m_2} \cdot \cdots \cdot \pi_n^{m_n},$$

where $\pi_1, \pi_2, \ldots, \pi_n$ are the first n primes.

Example 1. The Gödel number of the symbol s is $GN(s) = 2^{23}$. (Don't confuse numeric symbols and Gödel numbers!)

Example 2. The Gödel number of the number 2, which in S is represented by $ss0$, is

$$GN(ss0) = 2^{23} \cdot 3^{23} \cdot 5^{21} = 3765727153080000000000000000000.$$

Example 3. Let p be the string in the formal language S defined by

$$p := \forall y\, \exists x\, (x = sy).$$

This string is a well-formed formula with no free variables, *i.e.* a statement/proposition in S. It expresses Peano's second axiom, that every number has an immediate successor. The numeric symbols of the individual symbols in this string are $11, 31, 13, 29, 17, 29, 15, 23, 31, 19$. Hence the Gödel number of p is

$$GN(p) = 2^{11} \cdot 3^{31} \cdot 5^{13} \cdot 7^{29} \cdot 11^{17} \cdot 13^{29} \cdot 17^{15} \cdot 19^{23} \cdot 23^{31} \cdot 29^{19}.$$

Exercise 1. Find the Gödel number of the string $q := \forall x\big(\neg(0 = sx)\big)$. (Note that q is Peano's third axiom.) You may leave $GN(q)$ in factored form!

Exercise 2. Express Peano's fourth axiom as a statement r in S and find $GN(r)$ (in factored form).

Exercise 3. Let $T[x]$ be the arithmetical predicate, "x is an odd number." Express this predicate in the formal language S as a formula with one free variable, and determine its Gödel number (in factored form).

A theorem in the formal language S is a statement that can be proved from the axioms using the logical rules of inference, which are displayed in Table 2. A proof of a statement p is a sequence of statements, the last of which is p. For example, we can use universal instantiation to give a proof of the statement $0 \neq 1$:

$$q := \forall x(\neg(0 = sx));$$
$$z := \neg(0 = s0).$$

TABLE 3. Rules of Inference

Name	Rule of Inference
modus ponens	$p \implies q$ $\underline{p \qquad\quad}$ $\therefore q$
modus tollens	$p \implies q$ $\underline{\neg q \qquad\quad}$ $\therefore \neg p$
addition	$\underline{p \qquad\quad}$ $\therefore p \vee q$
simplification	$\underline{p \wedge q \quad}$ $\therefore p$
conjunction	p $\underline{q \qquad\quad}$ $\therefore p \wedge q$
hypothetical syllogism	$p \implies q$ $\underline{q \implies r}$ $\therefore p \implies r$
disjunctive syllogism	$p \vee q$ $\underline{\neg p \qquad\quad}$ $\therefore q$
universal instantiation	$\underline{\forall x\,(\phi(x))}$ $\therefore \phi(a)$ for any a in the universe of discourse
existential generalization	$\underline{\phi(a) \text{ for some } a \text{ in the universe of discourse}}$ $\therefore \exists x\,(\phi(x))$

We can extend Gödel numbering to sequences of strings. If $p_1, p_2, \dots p_n$ are strings in S with Gödel numbers $GN(p_1), GN(p_2), \dots, GN(p_n)$, then the Gödel number of the sequence $\sigma = (p_1, p_2, \dots p_n)$ is

$$GN(\sigma) = \pi_1^{GN(p_1)} \cdot \pi_2^{GN(p_2)} \cdot \dots \cdot \pi_n^{GN(p_n)},$$

where $\pi_1, \pi_2, \dots, \pi_n$ are the first n primes. For example, the Gödel number of the proof above that $0 \neq 1$ is

$$GN(p, z) = 2^{GN(q)} \cdot 3^{GN(z)}.$$

We can now assign a Gödel number to any symbol, string or sequence of strings in the formal system S. Conversely, can we determine if a given number is the Gödel number of some symbol, string or sequence of strings?

Exercise 4. Show that all Gödel numbers are even, but the converse is false.

Exercise 5. The number 1.4348907×10^7 is a Gödel number; of what?

2. Translating Meta-Mathematics into Arithmetic

The point of Gödel numbering is that by representing symbols, strings and sequences of strings in S with numbers, we can translate meta-mathematical sentences into purely arithmetical ones.

Exercise 6. Show that a string p begins with "$\neg$" if and only if $GN(p)$ is divisible by 2, but not by 4.

This exercise illustrates how a meta-mathematical sentence - a sentence about the formal system S - translates into an arithmetical sentence via Gödel numbering. In this case, the meta-mathematical sentence is a syntactical predicate, namely,

$$\phi(p) := p \text{ begins with "}\neg\text{".}$$

The corresponding arithmetical sentence is the arithmetical predicate

$$T[x] := x \text{ is divisible by 2, but not by 4.}$$

Exercise 7. Express $T[x]$ in the formal language S.

The biconditional statement in Exercise 6 could be expressed as:

$$\phi(p) \text{ if and only if } T[GN(p)].$$

Gödel showed that virtually all meta-mathematical sentences - namely the 'primitive recursive' ones - can be translated into arithmetical sentences via Gödel numbering. (A precise definition of 'primitive recursive' is beyond the scope of this worksheet.) For example, the meta-mathematical predicate "p is a well-formed formula," describing a syntactic property of a string in S, is a primitive recursive sentence, and so can be translated into a purely arithmetical predicate $Wff[GN(p)]$, describing an arithmetical property of the Gödel number of that string in $\mathbb{N}$. This arithmetic predicate will be very complicated, and we may well ask if we will be able to determine for any given (huge) number n whether it has this very complicated property. Our assumption that S is powerful enough to represent arithmetic means that we will certainly be able to *express* an arithmetical statement such as $Wff[n]$ in the language S. Importantly, Gödel showed that for any primitive recursive predicate, we will also be able to *prove* within S whether a specific number n does or does not satisfy the predicate.

Recall that the proof of a statement p in S in a sequence of statements, σ, the last of which is p. We can formulate this as a meta-mathematical binary relation:

$$\psi(p, \sigma) := \text{ statement } p \text{ is proved by sequence } \sigma.$$

Gödel showed that this binary relation is primitive recursive. Hence, the meta-mathematical relation ψ between the statement p and the sequence of strings σ can be represented by a purely arithmetical relation Prf between the Gödel numbers $GN(p)$ and $GN(\sigma)$:

(1) $$\psi(p, \sigma) \text{ if and only if } Prf\big[GN(p), GN(\sigma)\big].$$

Moreover, we will be able to prove within S whether or not any two specific numbers satisfy this arithmetical relation. From this arithmetical relation, we define the arithmetical predicate

$$Pr[x] := \exists y \Big(Prf[x, y]\Big).$$

In other words, for a given statement p in S,

(2) $$p \text{ is provable in } S \text{ if and only if } Pr[GN(p)].$$

The meta-mathematical property that a given statement is provable in S (*i.e.* is a theorem in S) can be translated into a purely arithmetical property of its Gödel number!

Moreover, since this property is primitive recursive, we have the following result.

Lemma 1. For any statement p in S, p is provable in S if and only if $Pr[GN(p)]$ is provable in S.

3. Diagonalization

Definition. Let $R = R[x]$ be an arithmetical predicate with Gödel number n. The *diagonalization* of R is the statement obtained by substituting the number n in for the free variable x in R:

$$R[n] = R\Big[GN\big(R[x]\big)\Big].$$

As with any statement, we can ask whether or not the diagonalization $R[n]$ is provable in S. In fact, we are more interested in when the diagonalization is **not** provable in S. Gödel showed that the meta-mathematical predicate "R is a predicate whose diagonalization is not provable in S" is primitive recursive. Thus, it can be translated into a purely arithmetical predicate U.

(3) $$R \text{ is a predicate whose diagonalization is not provable in } S \text{ if and only if } U\Big[GN\big(R[x]\big)\Big].$$

Put another way, the numbers that satisfy the arithmetical predicate U are precisely the Gödel numbers of arithmetical predicates whose diagonalizations are not provable in S.

Since U is an arithmetical predicate, it also has a diagonalization, which is called G: that is,

$$G := U\Big[GN\big(U[x]\big)\Big].$$

Theorem 1. *The biconditional statement "$G \iff \neg Pr[GN(G)]$" is provable in S.*

Proof. First assume that we assert the statement G. By definition, we are asserting $U\Big[GN\big(U[x]\big)\Big]$, which in turns asserts that the number $GN\big(U[x]\big)$ satisfies predicate U. By (3), this means that U is a predicate whose diagonalization is not provable in S. Since G is the diagonalization of predicate U, this means that G is not provable in S. By (2), this is asserting that the Gödel number of G does not satisfy the arithmetical predicate Pr, which is precisely $\neg Pr\Big[GN(G)\Big]$.

Conversely, assume that we assert $\neg Pr\big[GN(G)\big]$. This is asserting that the number $GN(G)$ does not have the arithmetical property of being the Gödel number of a statement that is provable in S. Since G is certainly a statement, this must mean that G is a statement that is not provable in S. By definition, G is the diagonalization of predicate U. Thus we are asserting that U is an arithmetical predicate whose diagonalization is not provable in S. By (3), the Gödel number of $U[x]$ must satisfy predicate U, which is precisely the assertion $U\Big[GN\big(U[x]\big)\Big]$, which is G. □

Note that this theorem is neutral on whether G is a 'true' statement; it's about provability, not truth. It merely asserts that if by starting with the axioms and using only the rules of logical inference we can get to statement G in S, then we can also get to statement $\neg Pr[GN(G)]$. Similarly, if by starting with the axioms and using only the rules of logical inference, we can get to $\neg Pr[GN(G)]$, then we can also get to G.

4. Consistency and Completeness

Definition. Let S be a formal system. Then S is

(1) *consistent* if and only if there is no statement p in S such that both p and $\neg p$ are provable in S;

(2) *complete* if and only if for every statement p in S, either p or $\neg p$ is provable in S.

We can express the definition of consistency in the formal language S with the statement

$$\neg\exists p\Big(Pr[GN(p)] \wedge Pr[GN(\neg p)]\Big),$$

or with the logically equivalent statement

$$\forall p\,\Big(\neg Pr[GN(p)] \vee \neg Pr[GN(\neg p)]\Big).$$

(Note here that we are quantifying over statements, *i.e.* well-formed formulas with no free variables.)

Exercise 8. Express the definition of completeness in the formal language S.

Now, if S is both consistent and complete, then for every statement p in S, exactly one of p or $\neg p$ is provable in S. This is obviously the gold standard for formal systems! Gödel showed that no formal system S that is powerful enough to represent Peano arithmetic can meet this gold standard.

GÖDEL'S FIRST INCOMPLETENESS THEOREM

If S is consistent, then S is incomplete.

Proof. We will show that if S is consistent, then neither G nor $\neg G$ is provable in S.

Claim 1. If S is consistent, then G is not provable in S.

Proof of Claim 1. We use proof by contrapositive (that is, *modus tollens*). Assume

$$G \text{ is provable in } S.$$

By Theorem 1, $G \iff \neg Pr[GN(G)]$ is provable in S, so by the inference rule of simplification,

$$G \implies \neg Pr[GN(G)] \text{ is provable in } S.$$

By *modus ponens*,

$$\neg Pr[GN(G)] \text{ is provable in } S.$$

However, we assumed G is provable in S and so by Lemma 1,

$$Pr[GN(G)] \text{ is provable in } S.$$

Since $\neg Pr[GN(G)]$ and $Pr[GN(G)]$ are both provable in S, by definition S is inconsistent.

Claim 2. If S is consistent, then $\neg G$ is not provable in S.

Proof of Claim 2. Assume that

$$\neg G \text{ is provable in } S.$$

By Theorem 1, $G \iff \neg Pr[GN(G)]$ is provable in S, so by simplification again,

$$\neg Pr[GN(G)] \implies G \text{ is provable in } S.$$

By *modus tollens*,

$$Pr[GN(G)] \text{ is provable in } S.$$

By Lemma 1, G is provable in S. Since both $\neg G$ and G are provable in S, by definition S is inconsistent. □

Gödel's Second Incompleteness Theorem states that a consistent formal system powerful enough to express Peano arithmetic cannot prove its own consistency. This does *not* mean that such a system is inconsistent, or that its consistency can never be proved; it only means that to prove the consistency of such a formal system, we need another (stronger) formal system.

To talk about a formal system S proving its own consistency, we must first formulate a statement in S that expresses the proposition that S is consistent. For this, we need to take a detour.

In classical logic, a contradiction is defined to be a compound statement (*i.e.* a statement made up of other constituent statements, combined using logical connectives) such that the final column in its truth table consists entirely of F's. In other words, the truth value of a contradiction is always 'false', no matter what the truth values of the constituent statements are. Clearly, for any statement p, the compound statement $p \wedge \neg p$ fits this definition of a contradiction. Now, if statement r is a contradiction, then for any statement q, the truth table for implication shows that $r \implies q$ is always true. Hence, if we can prove r, then we can prove any q.

The formalists, however, wanted to replace notions of truth and falsity with provability and non-provability. To replace the classical notion of contradiction, we introduce into our formal system a new symbol, which we will denote by $\perp$ (sometimes called the *absurdity constant*), which has the status of a formula with no free variables (*i.e.* a statement). We also introduce two new rules of inference associated with $\perp$:

$$\frac{p \wedge \neg p}{\therefore \perp} \qquad\qquad \frac{\perp}{\therefore \forall q(q)}$$

Exercise 9. Using the rules of inference for $\perp$ and in Table 2, prove that if a formal system S is inconsistent, then every statement q in S is provable in S. (This is sometimes called the *Principle of Explosion.*) [1]

The Principle of Explosion implies that if S is inconsistent, then for any statement p, *both* p and $\neg p$ are provable in S. Conversely, if *every* statement p in S has the property that both p and $\neg p$ are provable in S, then in particular *there exists* a statement p in S such that both p and $\neg p$ are provable in S (because universal quantification is stronger than existential quantification). Hence we have proved both directions of the biconditional:

$$S \text{ is inconsistent } \iff \forall p \Big(Pr[GN(p)] \wedge Pr[GN(\neg p)] \Big).$$

Taking the contrapositive yields:

$$S \text{ is consistent } \iff \exists p \Big(\neg Pr[GN(p)] \vee \neg Pr[GN(\neg p)] \Big).$$

In the first section, we showed that $z := \neg(0 = s0)$ (which is the expression in S of the arithmetical statement $0 \neq 1$) is provable in S. If S is consistent, then by definition $\neg z := 0 = s0$ is not provable in S. Conversely, if we can demonstrate that $\neg z$ is not provable in S, then by the above, S is consistent. We let

$$Con := \neg Pr[GN(0 = s0)].$$

Then Con is a statement in S that expresses the meta-mathematical statement that S is consistent.

[1] It is possible to construct a formal system without the Principle of Explosion. Such systems are called *paraconsistent*, and have been studied since the mid-twentieth century. The indirect method of proof by contradiction does not in general hold in such systems. Some argue that paraconsistent systems are an extension to classical systems, the same way as non-Euclidean geometry extends Euclidean geometry.

GÖDEL'S SECOND INCOMPLETENESS THEOREM.

If S is consistent, then Con is not provable in S.

Proof. Again, we use proof by contrapositive. Assume

$$Con \text{ is provable in } S.$$

The proof of Claim 1 in the proof of Gödel's First Incompleteness Theorem can be formalized in S, so

$$Con \Longrightarrow \neg Pr[GN(G)] \text{ is provable in } S.$$

By *modus ponens*,

$$\neg Pr[GN(G)] \text{ is provable in } S.$$

By Theorem 1, $G \Longleftrightarrow \neg Pr[GN(G)] \equiv (G \Longrightarrow \neg Pr[GN(G)]) \wedge (G \Longleftarrow \neg Pr[GN(G)])$ is provable in S. By simplification,

$$\neg Pr[GN(G)] \Longrightarrow G \text{ is provable in } S.$$

By *modus ponens* again,

$$GG \Longleftrightarrow \neg Pr[GN(G)]$$

By the contrapositive of Claim 1, S is inconsistent. $\square$

---

5. References

(1) Torkel Franzén, *Gödel's Theorem: An Incomplete Guide to its Use and Abuse*, A. K. Peters Ltd, 2005.

(2) Ernest Nagel and James Newman, *Gödel's Proof, Revised Edition*, New York University Press, 2001.

(3) Peter Smith, *An Introduction to Gödel's Theorems*, Cambridge University Press, 2007.

(4) Zach Weber, Inconsistent Mathematics, The Internet Encyclopedia of Philosophy ISSN 2161-0002, http://www.iep.utm.edu/math-inc, retrieved June 2016.

(5) Robert S. Wolf, *A Tour through Mathematical Logic*, Mathematical Association of America, 2005.

Thanks also to Professor Russell Marcus for helpful feedback on this worksheet.